Melina Freitag Patrick Kürschner

Tuned preconditioners for inexact two-sided inverse and Rayleigh quotient iteration

Max Planck Institute Magdeburg Preprints

MPIMD/13-04 **April 23, 2013**

Impressum:

Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg

Publisher:
Max Planck Institute for Dynamics of Complex Technical Systems

Address:
Max Planck Institute for Dynamics of Complex Technical Systems
Sandtorstr. 1
39106 Magdeburg

www.mpi-magdeburg.mpg.de/preprints

Tuned preconditioners for inexact two-sided inverse and Rayleigh quotient iteration

Melina A. Freitag* Patrick Kürschner†

April 23, 2013

Abstract

Convergence results are provided for inexact two-sided inverse and Rayleigh quotient iteration, which extend the previously established results to the generalized eigenproblem, and inexact solves with a decreasing solve tolerance. Moreover, the simultaneous solution of the forward and adjoint problem arising in two-sided methods is considered and the successful tuning strategy for preconditioners is extended to two-sided methods. Furthermore, it is shown that inexact two-sided Rayleigh quotient iteration and the inexact two-sided Jacobi-Davidson method (without subspace expansion) applied to the generalized preconditioned eigenvalue problem are equivalent when a certain number of steps of a Petrov-Galerkin-Krylov method is used, and when this specific tuning strategy is applied.

1 Motivation

Our aim is to find solutions of the two-sided generalized eigenvalue problem

$$Ax = \lambda Mx, \quad A^H y = \overline{\lambda} M^H y, \tag{1}$$

where A and $M \in \mathbb{C}^{n\times n}$ are assumed to be large and sparse matrices, and the nonzero vectors $x \in \mathbb{C}^n$ and $y \in \mathbb{C}^n$ are the right and left eigenvectors corresponding to the eigenvalue $\lambda \in \mathbb{C}^n$. For now we assume that the sought eigenvalues of A, M are finite and simple, such that $y^H Mx \neq 0$ is satisfied for the corresponding eigenvectors. If $M = I$ the finite condition number of a simple eigenvalue λ is given by $\kappa(\lambda) = |y^H x|^{-1}$ if $\|x\| = \|y\| = 1$. If only the right (or the left) eigenvectors are sought, (one-sided)

*Department of Mathematical Sciences, University of Bath, Claverton Down, BA2 7AY, United Kingdom, m.freitag@maths.bath.ac.uk.

†Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstraße 1, 39106 Magdeburg, Germany, kuerschner@mpi-magdeburg.mpg.de, corresponding author.

inverse or Rayleigh quotient iteration (RQI) provide basic methods for this purpose. In general (one-sided) inverse iteration converges linearly and RQI achieves quadratic convergence [21], in the normal case ($x = y$) RQI converges even cubically [20]. If inexact solves are used with a fixed solve tolerance then the order of convergence is reduced by one. If a decreasing tolerance proportional to the eigenvalue residual is chosen for the inexact solve, then the same convergence order as for exact solves can be recovered (see, for example [33, 5, 8, 7, 10]). It is also known that (one-sided) accelerated Rayleigh quotient iteration (Rayleigh quotient iteration with subspace expansion) is equivalent to the Jacobi-Davidson method [32] - if all linear systems are solved either exactly or, for Hermitian problems, by a certain number of steps of the conjugate gradient method [29]. This result has been extended to preconditioned non-Hermitian problems in [10], when a special preconditioner is used.

For this paper we are interested in the two-sided version of inverse iteration, Rayleigh quotient iteration and the related two-sided Jacobi-Davidson method [17, 34]. In [17] it was shown that, for the standard eigenproblem ($M = I$), two-sided Rayleigh quotient iteration converges locally cubically for exact solves and locally quadratically for inexact solves. We review these results in this paper, extend them to the generalized eigenvalue problem, and show that for inexact solves we can recover the convergence rate of the exact algorithms if we choose a decreasing solve tolerance, that is, linear convergence for inexact two-sided inverse iteration and cubic convergence for inexact two-sided Rayleigh quotient iteration.

Another important consideration independent of the convergence rate of the outer iteration is the choice of the preconditioner for the inner iteration. For the standard (one-sided) eigenproblem it has been shown that a "tuned" preconditioner, a rank-1 modification of the standard preconditioner, reduces the number of iterations for the inner solve considerably [8, 7, 11], a result that has been extended to inverse subspace iteration in [23, 38]. In this article we extend the result to two-sided inverse iteration and Rayleigh quotient iteration, where, due to the structure and the simultaneous solution of a forward and adjoint linear system a rank-2 modification of the standard preconditioner is necessary for an efficient tuning strategy.

In [17] it was shown that for the standard eigenproblem $M = I$ the simplified two-sided Jacobi-Davidson method (Jacobi-Davidson without subspace expansion) is equivalent to two-sided Rayleigh quotient iteration, when all pairs of linear equations are solved by a certain number of steps of a Petrov-Galerkin-Krylov method in each outer iteration. In this paper we extend this equivalence result to the generalized eigenproblem and to preconditioned solves, by using a special preconditioner and extensions of the techniques from [11, 10] for the one-sided algorithms.

In Section 2 we review and extend convergence results for two-sided inverse and Rayleigh quotient iteration. Section 3 contains preconditioning strategies for the inner iteration of the two-sided methods. In Section 4 we show the equivalence of two-sided Rayleigh quotient iteration and the Jacobi-Davidson method, when certain preconditioners are used. Section 5 supports our theory with numerical examples.

Algorithm 1: Two-sided inverse iteration (TII) and two-sided RQI (TRQI)

Input : Matrices A, M, initial vectors $u_1,\ v_1$, $v_1^H M u_1 \neq 0$.
Output: Approximate eigentriple $(\lambda_{k_{\max}},\ u_{k_{\max}},\ v_{k_{\max}})$.

1 $u_1 = u_1/\|u_1\|,\ v_1 = v_1/\|v_1\|$;
2 **for** $k = 1, 2, \ldots$ **do**
3 Set $\lambda_k = \rho(u_k,\ v_k)$;
4 Test for convergence;
5 Choose shift θ_k ;
6 Solve $(A - \theta_k M)u_{k+1} = Mu_k$ for u_{k+1} and normalize ;
7 Solve $(A - \theta_k M)^H v_{k+1} = M^H v_k$ for v_{k+1} and normalize ;

Notation used in this paper $\mathbb{R}$ and $\mathbb{C}$ denote the real and complex numbers and $\mathbb{R}^{n\times m}$, $\mathbb{C}^{n\times m}$ are $n \times m$ real and complex matrices, respectively. By $\overline{X}$ we denote the complex conjugate of a complex quantity X. The absolute value of $\xi \in \mathbb{C}$ is denoted by $|\xi|$. We use A^T and $A^H = \overline{A}^T$ for the transpose and, respectively, complex conjugate transpose of real and complex matrices. The inverse of a regular matrix A is denoted by A^{-1}. Expressions of the form $x = A^{-1}b$ should always be understood as solving the linear system of equations $Ax = b$ for x. The identity matrix is indicated by I. If not stated otherwise $\|\cdot\|$ is the Euclidean vector, or subordinate matrix norm. The expression $x^\perp$ stands for the orthogonal complement $\{z \in \mathbb{C}^n \backslash \{0\} \ : \ z \perp x\}$ of $x \in \mathbb{C}^n$.

2 Convergence theory for two-sided inverse iteration and RQI

The two-sided inverse iteration (TII) is illustrated in Algorithm 1 and requires a sufficiently good eigenvalue approximation $\theta \approx \lambda$ which is used a fixed shift. The two-sided RQI (TRQI) was originally proposed in [19] and is obtained by choosing the current shift θ_k in step 5 as the *two-sided generalized Rayleigh quotient*

$$\rho(u_k,\ v_k) := \frac{v_k^H A u_k}{v_k^H M u_k}, \tag{2}$$

of the previous iterates, where u_k and v_k are approximate right and left eigenvectors. Another early occurrence of this two-sided iteration can be found in [16, Section 13]. The main computational effort is done in the steps 6 and 7, where two linear systems with adjoint coefficient matrices have to be solved in each iteration of Algorithm 1 for obtaining new right and left eigenvector approximations. Throughout the rest of the paper we refer to the linear system in step 6 for u_{k+1} as *forward linear system*, and, respectively, the one in step 7 for v_{k+1} as *adjoint linear system*. Moreover, subscripts k always denote quantities of the kth iteration of Algorithm 1. In this section we review existing convergence results on the exact and inexact methods, where we adapt

the notation given in [17]. Note that here we consider the generalized eigenproblem as opposed to the standard eigenproblem $M = I$ in [17]. A result on inexact TRQI from [17] is then extended to the generalized eigenproblem and to inexact solves with decreasing solve tolerance.

2.1 Review of the convergence of the exact methods

For investigating the convergence rate of the methods, we assume that they converge, i.e., $u_k \to x$ and $v_k \to y$ and hence, $\theta_k \to \lambda$ as $k \to \infty$. Then it is possible to write

$$u_k = \alpha_k(x + \delta_k d_k) \quad \text{and} \quad v_k = \beta_k(y + \epsilon_k e_k), \tag{3}$$

where $\delta_k, \epsilon_k \geq 0$, $u_k,\ v_k,\ d_k,\ e_k,\ x,\ y$ are unit vectors with $d_k \perp M^H y$, $e_k \perp Mx$ and α_k, β_k are non-zero normalization constants. Moreover, the matrices $A - \mu M$, $A^H - \overline{\mu} M^H$ are surjective mappings from $(M^H y)^\perp$ to $y^\perp$ and $(Mx)^\perp$ to $x^\perp$, respectively, for all $\mu \in \mathbb{C}$ in a sufficiently small vicinity of λ. Thus, there exist a $\psi \geq 0$ such that

$$\|(A - \mu M)s\| \geq \psi \|s\| \quad \text{and} \quad \|(A - \mu M)^H t\| \geq \psi \|t\| \tag{4}$$

for all μ in a neighborhood of λ and $s \in (M^H y)^\perp$, $t \in (Mx)^\perp$ (cf. [25]). The following result extends the work from [17] to the generalized eigenproblem.

Theorem 1 (Convergence of exact two-sided II and RQI)**.** *Under the above assumptions, let u_k, v_k be given as in (3). Then the following convergence results hold.*

1. *For TII, let $\theta_k \equiv \theta$ be approximation to the simple eigenvalue λ with $|\theta - \lambda| = \nu$. Then Algorithm 1 converges linearly, i.e.,*

$$\delta_{k+1} \leq \gamma^{TII} \delta_k + \text{h.o.t.}, \quad \epsilon_{k+1} \leq \gamma^{TII} \epsilon_k + \text{h.o.t.}, \tag{5}$$

   where $\gamma^{TII} := \nu \|M\| / \psi^{TII}$.

2. *TRQI achieves locally cubic convergence, i.e.,*

$$\delta_{k+1} \leq \gamma^{TRQI} \delta_k^2 \epsilon_k + \text{h.o.t.} \quad \textit{and,} \quad \epsilon_{k+1} \leq \gamma^{TRQI} \delta_k \epsilon_k^2 + \text{h.o.t.} \tag{6}$$

   with $\gamma^{TRQI} := \omega(\lambda)\|M\|\|A - \lambda M\| / \psi^{TRQI}$ and $\omega(\lambda) := |y^H M x|^{-1}$.

The constants $\psi^{TII},\ \psi^{TRQI} \geq 0$ originate from (4) when using a constant or Rayleigh quotient shift and h.o.t. *stands for higher-order terms in δ_k and ϵ_k.*

Proof. As in the proof of [17, Theorem 3.1] for TRQI, which is a slight generalization of the original convergence proof in [20, p. 689], one can find nonzero $\alpha_{k+1},\ \beta_{k+1}$ such that

$$\begin{aligned} u_{k+1} &= \alpha_{k+1} \left(x + \delta_k (\lambda - \theta_k) \hat{d}_k \right), \\ v_{k+1} &= \beta_{k+1} \left(y + \epsilon_k (\overline{\lambda - \theta_k}) \hat{e}_k \right) \end{aligned} \tag{7}$$

with $\hat{d}_k$,$\hat{e}_k$ from $(A-\theta_k M)\hat{d}_k = Md_k$ and $(A-\theta_k M)^H\hat{e}_k = M^H e_k$. Applying (4) we have

$$\|\hat{d}_k\| \leq \frac{\|(A-\mu M)\hat{d}_k\|}{\psi} = \frac{\|Md_k\|}{\psi} \leq \frac{\|M\|}{\psi},$$
$$\|\hat{e}_k\| \leq \frac{\|(A-\mu M)^H\hat{e}_k\|}{\psi} = \frac{\|M^H e_k\|}{\psi} \leq \frac{\|M\|}{\psi},$$

and the result for exact TII in (5) follows. Basic calculations yield

$$|\lambda-\theta_k| = \left|\frac{\delta_k\epsilon_k e_k^H(A-\lambda M)d_k}{y^H Mx + \delta_k\epsilon_k e_k^H Md_k}\right| \leq \omega(\lambda)\delta_k\epsilon_k|e_k^H(A-\lambda M)d_k| + \text{h.o.t.} \tag{8}$$

which reveals that (2) can (like the one-sided Rayleigh quotient) be seen as a quadratically accurate eigenvalue approximation. Replacing ν in the TII result with this estimate gives the desired result (6). □

A different approach showing the cubic rate of convergence can be found in [1].

2.2 Convergence under inexact solves

We are especially interested in the case when the involved linear systems in steps 6 and 7 are solved inexactly, i.e. for instance

$$\begin{aligned}\|(A-\theta_k M)u_{k+1} - Mu_k\| &\leq \xi_k^{\mathrm{R}}\|Mu_k\| < 1,\\ \|(A-\theta_k M)^H v_{k+1} - M^H v_k\| &\leq \xi_k^{\mathrm{L}}\|M^H v_k\| < 1,\end{aligned} \tag{9}$$

where the constants $\xi_k^{\mathrm{R}}, \xi_k^{\mathrm{L}}$ define the accuracy to which the linear systems are solved in the kth iteration of Algorithm 1. This inexact solution is usually carried out using Krylov subspace methods for unsymmetric linear systems, e.g., GMRES, BiCG, BiCGstab, BiCGStab(ℓ), QMR, IDR(s) [4, 26] to name at least a few. From now on, the method used to solve the linear system is referred to as inner solver and its iterations are called inner iterations. The iterations of eigenvalue method TII or TRQI (Algorithm 1) are hence referred to as the outer iterations.

For investigating the convergence of inexact two-sided inverse and Rayleigh quotient iteration, the following statement is helpful which immediately follows from (9):

$$(A-\theta_k M)u_{k+1} = M(u_k + \tilde{\xi}_k^{\mathrm{R}}\|Mu_k\|f_k), \tag{10}$$
$$(A-\theta_k M)^H v_{k+1} = M^H(v_k + \tilde{\xi}_k^{\mathrm{L}}\|M^H v_k\|g_k), \tag{11}$$

where $0 \leq \|Mf_k\|\tilde{\xi}_k^{\mathrm{R}} \leq \xi_k^{\mathrm{R}}$, $0 \leq \|M^H g_k\|\tilde{\xi}_k^{\mathrm{L}} \leq \xi_k^{\mathrm{L}}$ and f_k, g_k are unit vectors. The following Theorem extends [17, Theorem 5.2] to the generalized eigenproblem.

Theorem 2 (Convergence with fixed inner accuracies [17, Theorem 5.2]). *Let* $\xi_k^R \equiv \xi^R$, $\xi_k^L \equiv \xi^L$ $\forall k$ *in* (9) *with* $\max(\xi^R\|Mu_k\|, \xi^L\|M^H v_k\|)\omega(\lambda) < 1$ *and define*

$$\zeta_k^R := \frac{\xi^R\omega(\lambda)\|Mu_k\|}{1-\xi^R\omega(\lambda)\|Mu_k\|}, \quad \zeta_k^L := \frac{\xi^L\omega(\lambda)\|M^H v_k\|}{1-\xi^L\omega(\lambda)\|M^H v_k\|}.$$

1. *Then we have for one step of inexact two-sided inverse iteration*

$$\delta_{k+1} \leq \gamma^{TII}\zeta_k^R + \text{h.o.t.}, \quad \textit{and} \quad \epsilon_{k+1} \leq \gamma^{TII}\zeta_k^L + \text{h.o.t.}. \tag{12}$$

2. *For one step of inexact TRQI it holds*

$$\delta_{k+1} \leq \delta_k\epsilon_k\gamma^{TRQI}\zeta_k^R + \text{h.o.t.}, \quad \textit{and} \quad \epsilon_{k+1} \leq \delta_k\epsilon_k\gamma^{TRQI}\zeta_k^L + \text{h.o.t.}, \tag{13}$$

i.e., inexact TRQI with fixed inner tolerances converges locally quadratic.

The constants γ^{TII}, γ^{TRQI} are as in Theorem 1.

Proof. Exemplary, for the forward linear system, using (3) and decomposing f_k as

$$f_k = \frac{xy^HM}{y^HMx}f_k + \left(I - \frac{xy^HM}{y^HMx}\right)f_k,$$

we find

$$\begin{aligned} u_k + \tilde{\xi}_k^{\mathrm{R}}\|Mu_k\|f_k &= \left(\alpha_k + \tilde{\xi}_k^{\mathrm{R}}\|Mu_k\|\frac{y^HMf_k}{y^HMx}\right)x \\ &\quad + \alpha_k\delta_k d_k + \tilde{\xi}_k^{\mathrm{R}}\|Mu_k\|\left(I - \frac{xy^HM}{y^HMx}\right)f_k \\ &= \tilde{\alpha}_k\left(x + \frac{\tilde{\delta}_k}{\tilde{\alpha}_k}\tilde{d}_k\right), \quad \tilde{d}_k \perp M^Hy. \end{aligned}$$

Since $|\tilde{\alpha}_k| \geq |\alpha_k| - \tilde{\xi}_k^{\mathrm{R}}\|Mf_k\|\|Mu_k\|\omega(\lambda)$, $\tilde{\delta}_k \leq |\alpha_k|\delta_k + \tilde{\xi}_k^{\mathrm{R}}\|Mf_k\|\|Mu_k\|\omega(\lambda)$ and $|\alpha_k| = 1 + \text{h.o.t.}$ we obtain

$$\left|\frac{\tilde{\delta}_k}{\tilde{\alpha}_k}\right| = \zeta_k^{\mathrm{R}} + \text{h.o.t.},$$

using $\|Mf_k\|\tilde{\xi}_k^{\mathrm{R}} \leq \xi_k^{\mathrm{R}} \equiv \xi^{\mathrm{R}}$. The remainder of the proof is similar to the proof of Theorem 1. □

Theorem 2 shows that inexact TII can stagnate if the steps 6, 7 of Algorithm 1 are solved to a fixed accuracy, whereas inexact TRQI achieves quadratic convergence, as we would expect.

For one-sided methods it can be shown that, by using increasing inner accuracies, i.e. decreasing sequences

$$\xi_k^{\mathrm{R}} \leq \xi_{k-1}^{\mathrm{R}}, \quad \xi_k^{\mathrm{L}} \leq \xi_{k-1}^{\mathrm{L}},$$

the convergence rate of the exact methods can be reestablished in the inexact ones [8, 11, 36]. The next theorem shows that this can also be achieved for the two-sided methods by asking that ξ_k^{R} and ξ_k^{L} are proportional to the eigenvalue residual norms $\|r_{u_k}\|$ and $\|r_{v_k}\|$, respectively, which is similar to the one-sided case. Hence we can achieve the same convergence rates for inexact TII and inexact TRQI as for the exact versions of the algorithms, if a decreasing inner solve tolerance is used.

Theorem 3 (Convergence with decreasing inner tolerances)**.** *Let the assumptions of Theorem 2 hold but choose*

$$\xi_k^R = \eta_k^R \|r_{u_k}\| \quad \textit{and} \quad \xi_k^L = \eta_k^L \|r_{v_k}\|, \tag{14}$$

for some $\eta_k^R, \eta_k^L > 0$*, where* $r_{u_k} = (A - \theta_k M)u_k$ *and* $r_{v_k} = (A - \theta_k M)^H v_k$*. Then there exist* $\eta_k^R, \eta_k^L < 1$ *such that the following estimates hold.*

1. *For inexact TII:*

$$\delta_{k+1} < \gamma^{TII}\delta_k + \text{h.o.t.} \quad \textit{and,} \quad \epsilon_{k+1} \leq \gamma^{TII}\epsilon_k + \text{h.o.t.} \tag{15}$$

2. *For inexact TRQI:*

$$\delta_{k+1} < \gamma^{TRQI}\delta_k^2\epsilon_k + \text{h.o.t.}, \quad \textit{and} \quad \epsilon_{k+1} \leq \gamma^{TRQI}\epsilon_k^2\delta_k + \text{h.o.t.} \tag{16}$$

The constants γ^{TII}*,* γ^{TRQI} *are again the ones from Theorem 1.*

Proof. We proceed through the proof exemplary for the linear system for the right eigenvector approximation. The result for ϵ_{k+1} in the adjoint linear system for the left eigenvector approximations can be obtained similarly. Using (3), (8) and $|\alpha_k| = 1 + \text{h.o.t.}$ one finds

$$\begin{aligned} \|r_{u_k}\| = \|Au_k - \theta_k M u_k\| &= \|\alpha_k(\lambda - \theta_k)Mx - \alpha_k\delta_k(A - \theta_k M)d_k\| && (17)\\ &\leq \delta_k\|(A - \theta_k M)d_k\| + \text{h.o.t.}, && (18)\end{aligned}$$

and using (14) this can be plugged into the expression for ζ_k^{R} to obtain

$$\zeta_k^{\mathrm{R}} \leq \frac{\eta_k^{\mathrm{R}}\delta_k\|Mu_k\|\|(A - \theta_k M)d_k\|\omega(\lambda)}{1 - \eta_k^{\mathrm{R}}\delta_k\|Mu_k\|\|(A - \theta_k M)d_k\|\omega(\lambda)}. \tag{19}$$

The fraction on the right hand side of the above expression can be bounded from above by δ_k if

$$\eta_k^{\mathrm{R}} < \left((1 + \delta_k)\|Mu_k\|\|(A - \theta_k M)d_k\|\omega(\lambda)\right)^{-1}.$$

Assuming that this inequality holds, the result for inexact TII then follows from using (19) in (12). For the left eigenvector approximation one finds in a similar way that

$$\eta_k^{\mathrm{L}} < \left((1 + \epsilon_k)\|M^H v_k\|\|(A - \theta_k M)^H e_k\|\omega(\lambda)\right)^{-1}.$$

has to hold. The result for inexact TRQI follows as before from (8). □

One can conclude inexact TII and TRQI, converge at a similar speed as their exact versions if decreasing solve tolerances are used.

Remark 4. *Note that due to the large constants in any of the previous theorems, the actual convergence might be slower than stated for highly nonnormal problems. Hence, it might be more adequate to rephrase the statement of Theorem 3 in the sense that the convergence of the inexact methods can be improved by using increasing inner accuracies.*

Remark 5. *In order to implement the decreasing solve tolerance in practice one usually chooses* $\xi_k^R = \min\{\varphi_0^R, \|r_{u_k}\|\}$ *or* $\xi_k^R = \min\{\varphi_1^R, \varphi_2^R\|r_{u_k}\|\}$*, and* $\xi_k^L = \min\{\varphi_0^L, \|r_{v_k}\|\}$ *or* $\xi_k^L = \min\{\varphi_1^L, \varphi_2^L\|r_{v_k}\|\}$*, where* $\varphi_i^R < 1$*,* $\varphi_i^L < 1$*,* $i = 0, 1, 2$*, are sufficiently small constants.*

Inspired by TRQI a two-sided Jacobi-Davidson method with bi-orthogonal basis vectors can be designed [34, 17] (see Section 4). It is well known that a simplified version of the Jacobi-Davidson method (a method without subspace expansion) is equivalent to Rayleigh quotient iteration if the linear systems are solved exactly. Hence, the same convergence rates as stated in Theorem 1 can be expected (see also [17, Theorem 4.1]).

For the inexact versions of both algorithms the equivalence between the methods does not hold and hence the convergence theories cannot be carried over. In fact, in [17, Theorem 5.3]) it was shown that two-sided Jacobi-Davidson achieves locally linear convergence whereas TRQI converges locally quadratically (see Theorem 2) for a fixed solve tolerance. However, as also noted in [17], this does not mean that inexact two-sided JD has a worse behavior than inexact TRQI.

In Section 4 we will show that both methods are in fact still equivalent for inexact solves, if a certain preconditioner and type of linear system solver is used - and hence the convergence theories for both methods can be carried over.

3 The inner iteration and tuned preconditioners

A good preconditioner helps to speed up the computations in step 6 and 7 of Algorithm 1. In this section we investigate the behavior of preconditioned Krylov subspace methods used to solved the linear systems (9) inexactly. We distinguish between two strategies: On the one hand solving those two linear systems separately by two runs of a Krylov subspace method or, on the other hand, since the matrices in both linear systems are adjoint to each other, simultaneously by suitable Krylov subspace methods. The first case essentially amounts to follow the approaches for one-sided eigenvalue problems [10, 8, 9, 11]. There the concept of *tuning*, and more precisely *tuned preconditioners*, is introduced which significantly reduces the amount of work to achieve (9) in terms of the required inner iterations. The main contribution of this section is then to carry this *tuning* concept over to the simultaneous solution. In this section superscripts (i) refer to quantities related to the ith inner iteration, i.e. the ith iteration of the applied Krylov subspace methods.

3.1 Separate solution

If the systems in (9) are solved separately, a whole range of Krylov subspace methods are available. For illustration purposes GMRES is considered here. For the solution of the linear system $Cu_{k+1} = Mu_k$ or $C^H v_{k+1} = M^H v_k$, where $C = (A-\theta_k M)$, consider, exemplary, $Cx = b$. If the right hand side b is an eigenvector of C and we start GMRES (or in fact any Krylov method) with zero initial guess, then the method will converge within one step. For inexact inverse iteration or Rayleigh quotient iteration without a preconditioner and $M = I$, the right hand side b turns out to be an approximation to the eigenvector of C which becomes increasingly better as the outer iteration proceeds (see [29]) and hence an improved performance of the inner iteration is observed, even if the system becomes more and more singular. If (left or right) preconditioners are applied, this property gets lost and the right hand side b is generally far from a good approximation to the eigenvector of C. For the standard, one-sided methods tuned preconditioners have been proposed, which are rank-one updates of the standard preconditioners and force the right hand side of the linear system to be an approximation to the eigenvector of the system matrix, in fact, in the limit, if the outer iteration has converged, they are exact eigenvectors of the system matrix [10, 8, 9, 11, 37].

For the analysis here we concentrate on a fixed shift $\theta_k = \theta$ (that is TII) and the forward linear system, hence, in the unpreconditioned case $C = A-\theta M$ and $b = Mu_k$. The motivation for the adjoint linear system and the variable shift (that is TRQI) is similar. Moreover, we choose decreasing inner tolerances $\xi_k^R = \eta^R \|r_{u_k}\|$ in order to obtain convergence of the method, see Theorem 3. We also explain our strategy using a right preconditioner, but the theory extends to the case of a left preconditioner (see [11]).

Let the smallest eigenvalue c_1 of C be separated from the other $n-1$ eigenvalues. We can block-diagonalise C as

$$C = [v_{c_1} V_{C_2}] \begin{bmatrix} c_1 & 0 \\ 0 & C_2 \end{bmatrix} [v_{c_1} V_{C_2}]^{-1}, \tag{20}$$

where $\|v_{c_1}\| = 1$ and V_{C_2} has orthonormal columns. We have the following Lemma which explains the role the right hand side plays in the solution of the linear system $Cx = b$ when Krylov subspace methods are used. It follows directly from [23, Theorem 3.7],[38, Lemma 3.1].

Lemma 6. *[38, Lemma 3.1] Suppose the field of values* $W(C_2) = \{\frac{z^* C_2 z}{z^* z} : z \in \mathbb{C}^{n-1}, z \neq 0\}$ *or the* ε*-pseudospectrum* $\Lambda_\varepsilon = \{z \in \mathbb{C}^{n-1} : \|(zI - C_2)^{-1}\| > \varepsilon\}$ *is contained in a convex closed bounded set* E *in the convex plane with* $0 \notin E$*. Assume GMRES is used to solve* $Cx = b$*, where* $b \in \mathbb{C}^n$ *can be decomposed as* $b = v_{c_1} b_1 + V_{C_2} b_2$*, where* v_{c_1} *and* V_{C_2} *are given in (20) and* $b_2 \in \mathbb{C}^{n-1}$*,* $b_1 \neq 0$*. Let* $x^{(i)}$ *be the approximate solution to* $Cx = b$ *obtained after the* i*th GMRES iteration with* $x^{(0)} = 0$*. If*

$$i \geq 1 + D_a \left(D_b + \log \frac{\|b_2\|}{\xi \|b\|} \right), \tag{21}$$

then $\frac{\|b - Cx^{(i)}\|}{\|b\|} < \xi$*. The constants* D_a *and* D_b *depend on the spectrum of* C*.*

For details about D_a and D_b we refer the reader to [23, 11], for our analysis these constants have very little significance. We have the following Theorem for the number of inner iterations of the unpreconditioned algorithm (see [38, Theorem 3.2]) applied to the forward linear system.

Theorem 7. *Assume that (un)preconditioned GMRES is used to solve the linear system $Cu_{k+1} = Mu_k$, where $C = A - \theta M$ to the prescribed tolerance in (9), that is, $\|(A-\theta M)u_{k+1} - Mu_k\| \leq \xi_k^R \|Mu_k\| < 1$ and $\xi_k^R = \eta^R \|r_{u_k}\|$ from (14). Then under assumptions in Theorem 3 we have $u_k \to x$ linearly, and the lower bound on the GMRES iterations from Lemma 6 increases as the outer iteration proceeds.*

Proof. From (21) the lower bound on the GMRES iterations is

$$i_k \geq 1 + D_a \left(D_b + \log \frac{\|(b_2)_k\|}{\eta^R \|r_{u_k}\| \|Mu_k\|} \right). \tag{22}$$

Using (17) we have

$$\eta^R \|r_{u_k}\| \|Mu_k\| \leq \delta_k \|(A - \theta_k M) d_k\| \eta^R \|M\| + \text{h.o.t.},$$

and clearly the denominator in the second term of (22) converges to zero as $u_k \to x$ since $\delta_k \to 0$. The nominator $(b_2)_k$ is the component of $Mu_k = v_{c_1}(b_1)_k + V_{C_2}(b_2)_k$ which is not in the direction of the eigenvector v_{c_1} of C. As $Mu_k \to Mx$ which is not an eigenvector of $C = A - \theta M$ we have $\|(b_2)_k\| \to \|b_2\| \neq 0$, where $Mx = v_{c_1} b_1 + V_{C_2} b_2$. Hence, the expression on the right of (22) increases as the outer iteration k increases. □

Remark 8. *An equivalent result to Theorem 7 holds for the adjoint linear system in step 7 of Algorithm 1. Note that Theorem 7 only shows that a lower bound on i_k increases, there is no result on the behaviour of the actual iteration count. However, numerical results in [11] endorse that our theoretical findings on the iteration bound are indicative of the actual performance of the iterations.*

Now consider the (right) preconditioned systems

$$CP^{-1}\tilde{u}_{k+1} = Mu_k, \quad P^{-1}\tilde{u}_{k+1} = u_{k+1}, \tag{23}$$

where P is a preconditioner for $C = A - \theta M$. The results of Theorem 7 also hold for the preconditioned system (23), as, in general, the right hand side Mu_k is not an approximate eigenvector of CP^{-1}. (In the limit, Mx is not an eigenvector of CP^{-1}.)

A new type of preconditioner was considered in [8, 11]. Assume we have preconditioners $\mathbb{P}_k$ such that

$$\mathbb{P}_k u_k = Mu_k \quad \text{or} \quad \mathbb{P}_k u_k = Au_k. \tag{24}$$

In both cases the right hand side Mu_k is an approximate eigenvector of $C\mathbb{P}_k^{-1}$, since, using (3)

$$\begin{aligned} (A-\theta M)\mathbb{P}_k^{-1} Mu_k - (\lambda - \theta) Mu_k &= (A - \lambda M)u_k \\ &= \alpha_k \delta_k (A - \lambda M) d_k \end{aligned}$$

or, assuming $\lambda \neq 0$,

$$\begin{aligned}(A-\theta M)\mathbb{P}_k^{-1}Mu_k - \frac{\lambda-\theta}{\lambda}Mu_k &= \frac{A-\lambda M}{\lambda}u_k - (A-\theta M)\mathbb{P}_k^{-1}\frac{A-\lambda M}{\lambda}u_k \\ &= \alpha_k\delta_k\left(\frac{I-(A-\theta M)\mathbb{P}_k^{-1}}{\lambda}\right)(A-\lambda M)d_k,\end{aligned}$$

and, as $u_k \to x$ we have $\delta_k \to 0$ and Mu_k is an approximate eigenvector of $(A-\theta M)\mathbb{P}_k^{-1}$. For the choices of $\mathbb{P}_k$ in (24) we then have $\|(b_2)_k\| = D_c\delta_k$ in Theorem 7 and the lower bound on the iteration number does not increase. We observe this in the numerical examples in [11].

In order to satisfy the first condition in (24) we may choose

$$\mathbb{P}_k = I + (Mu_k - u_k)u_k^H \quad \text{or} \quad \mathbb{P}_k = P + (Mu_k - Pu_k)u_k^H, \tag{25}$$

depending on the availability of a preconditioner. Similar we may choose

$$\mathbb{P}_k = I + (Au_k - u_k)u_k^H \quad \text{or} \quad \mathbb{P}_k = P + (Au_k - Pu_k)u_k^H, \tag{26}$$

in order to satisfy the second condition in (24). Note that these choices are just rank-one updates of the identity or the preconditioner and can be efficiently implemented using the Sherman-Morrison formula [15].

Remark 9. *For the adjoint linear system we use a tuned preconditioner $\mathbb{Q}_k$ such that*

$$\mathbb{Q}_k v_k = M^H v_k \quad \text{or} \quad \mathbb{Q}_k v_k = A^H v_k. \tag{27}$$

instead of (24) in order to satisfy the property that the right hand side of the linear system is an approximation for the eigenvector of the system matrix. This can be achieved by

$$\mathbb{Q}_k = P^H + (M^H v_k - P^H v_k)v_k^H \tag{28}$$

$$\text{or} \quad \mathbb{Q}_k = P^H + (A^H v_k - P^H v_k)v_k^H. \tag{29}$$

Note that the strategy of tuning can in fact be applied to any Krylov method (that is, if the right hand side of the system is an approximation for the eigenvector of the system we obtain convergence in few steps), however, explicit bounds are only available for certain Krylov methods such as GMRES.

For steps 6 and 7 of Algorithm 1 it can be more efficient to solve both the forward and adjoint systems simultaneously, which we will consider in the next section. The concept of tuning for simultaneous solution is examined in Section 3.3.

3.2 Simultaneous solution

Since the coefficient matrices in both linear systems are adjoint to each other, they can be solved simultaneously to the desired accuracy in one run of Krylov subspace

methods which are based on the two-sided (non-symmetric) Lanczos process. There, for the solution of a linear system $Cx = b$, biorthogonal bases for the Krylov subspaces

$$\mathcal{K}_{\mathrm{R}}^{(i)}(C, s) = \operatorname{span}\left\{s, Cs, \ldots, C^{i-1}s\right\},$$
$$\mathcal{K}_{\mathrm{L}}^{(i)}(C^H, t) = \operatorname{span}\left\{t, C^H t, \ldots, (C^H)^{i-1}t\right\}$$

are build for some generating Krylov vectors s and t. Probably the most prominent methods belonging to this class are the bi-conjugate gradient (BiCG) [6] and Quasi Minimal Residual method (QMR) [13], which can also implicitly solve an adjoint system $C^H y = c$ if the generating vectors are chosen suitably, e.g. $s = b - Cx^{(0)}$, $t = c - C^H y^{(0)}$ for some initial guesses $x^{(0)}$, $y^{(0)}$. This is, however, only rarely exploited in practice [18, 14], but here we make explicit use of this property.

The two-sided Lanczos process is the generalization of standard Lanczos process for nonsymmetric matrices. Let the columns of $W^{(i)}$ and $Z^{(i)}$ span biorthonormal bases for $\mathcal{K}_{\mathrm{R}}^{(i)}$ and $\mathcal{K}_{\mathrm{L}}^{(i)}$, and assume for simplicity that zero initial vectors $x^{(0)}$, $y^{(0)}$ are used, then the approximate solutions after i steps of the two-sided Lanczos for linear systems [26, Algorithm 7.2] are given by

$$x^{(i)} = W^{(i)}(T^{(i)})^{-1}(Z^{(i)})^H b, \quad y^{(i)} = Z^{(i)}(T^{(i)})^{-H}(W^{(i)})^H c,$$

where $T^{(i)} := (Z^{(i)})^H C W^{(i)}$ is tridiagonal. In other words, the approximate solutions are constructed according to a Petrov-Galerkin projection of the linear system, and hence we call methods following this approach *Petrov-Galerkin methods* for solving the linear systems when not referring to a particular implementation.

Recursively updating an LU decomposition of $T^{(i)}$ for solving the small $i \times i$ linear systems leads to the short-recurrence formulation of BiCG [6] which is more economical in terms of the amount of work since only a small and constant number of vectors needs to be stored in each step to get $x^{(i)}$, $y^{(i)}$. In Algorithm 2 the basic preconditioned BiCG algorithm is illustrated. It is possible that Algorithm 2 breaks down. On the one hand when the underlying two-sided Lanczos process breaks down, i.e. $\gamma^{(i)} = 0$ but $s^{(i)}$, $g^{(i)}$ are nonzero vectors, which is referred to as *breakdown of the first kind* or *Lanczos breakdown*. On the other hand, it is possible that a pivotless LU decomposition of $T^{(i)}$ does exist which is called *breakdown of the second kind* or *pivot breakdown*. This happens in Algorithm 2 when $(q^{(i)})^H v^{(i)} = 0$ in step 8. Lanczos breakdowns can be handled by sophisticated look-ahead strategies (see, e.g., [12] and the references therein) which leads to rather complicated expressions where the short-recurrence property of BiCG is also lost to some extend. In the remainder we assume that these breakdowns do not occur. Breakdowns of the second kind can be dealt with, e.g., the composite step BiCG (CSBCG) [3] which is a rather simple rearrangement of Algorithm 2. Since these pivot breakdowns appear to be a special issue in the context of solving the adjoint linear systems of two-sided RQI, they are investigated in more depth in the next subsection.

Other less prominent methods which are also capable of a simultaneous solution of forward and adjoint linear system are GLSQR [14, 22, 27] as well as unsymmetric variants of MINRES and SYMMLQ [27], where other Krylov-like subspaces which are not related to the two-sided Lanczos process are used.

Algorithm 2: Preconditioned BiCG [4]

Input : $C,\ b,\ c$, initial guesses $x^{(0)}, y^{(0)}$, preconditioner $P \approx C$.
Output: $\tilde{x},\ \tilde{y}$ approximately solving $Cx = b$, $C^H y = c$.

1 $\gamma^{(-1)} = 1$, $p^{(0)} = q^{(0)} = 0$;
2 $f^{(0)} = b - Cx^{(0)}$, $g^{(0)} = c - C^H y^{(0)}$;
3 **for** $i = 1, 2, \ldots,$ **do**
4 Solve $Ps^{(i-1)} = f^{(i-1)}$, $P^H t^{(i-1)} = g^{(i-1)}$ for $s^{(i-1)}$, $t^{(i-1)}$.;
5 $\gamma^{(i-1)} = (g^{(i-1)})^H s^{(i-1)}$, $\beta^{(i-1)} = \frac{\gamma^{(i-1)}}{\gamma^{(i-2)}}$;
6 $p^{(i)} = s^{(i-1)} + \beta^{(i-1)} p^{(i-1)}$, $q^{(i)} = t^{(i-1)} + \overline{\beta^{(i-1)}} q^{(i-1)}$;
7 $v^{(i)} = Cp^{(i)}$, $w^{(i)} = C^H q^{(i)}$;
8 $\alpha^{(i)} = \frac{\gamma^{(i-1)}}{(q^{(i)})^H v^{(i)}}$;
9 $x^{(i)} = x^{(i-1)} + \alpha^{(i)} p^{(i)}$, $y^{(i)} = y^{(i-1)} + \overline{\alpha^{(i)}} q^{(i)}$;
10 $f^{(i)} = f^{(i-1)} - \alpha^{(i)} v^{(i)}$, $g^{(i)} = g^{(i-1)} - \overline{\alpha^{(i)}} w^{(i)}$;

3.2.1 Breakdowns of BiCG with two-sided RQI

Before discussing tuned preconditioners for the simultaneous solution, we shall have a look at the issue of breakdowns of BiCG within two-sided Rayleigh quotient iteration. Apply the above BiCG to the adjoint linear systems occurring in, say iteration k of TRQI, i.e.

$$C = A - \theta_k M,\ b = Mu_k,\ c = M^H v_k,$$

where $u_k,\ v_k$ are approximations to right and left eigenvectors of $A,\ M$, respectively.

Theorem 10 (Breakdown of BiCG within two-sided RQI)**.** *Consider Algorithm 2 applied within TRQI with $x^{(0)} = y^{(0)} = 0$. Then BiCG suffers from a breakdown of the second kind in the very first iteration if $M = P = I_n$ or $Pu_k = Mu_k$, $P^H v_k = M^H v_k$ holds.*

Proof. For $M = P = I_n$, the first BiCG iteration gives

$$p^{(1)} = s^{(0)} = f^{(0)} = u_k, \quad q^{(1)} = t^{(0)} = g^{(0)} = v_k.$$

With $v^{(1)} = (A - \theta_k I_n) p^{(1)}$ and θ_k being the two-sided Rayleigh quotient this yields

$$(q^{(1)})^H v^{(1)} = v_k^H (A - \theta_k I_n) u_k = v_k^H \left(A - \frac{v_k^H A u_k}{v_k^H u_k} I_n \right) v_k = 0$$

such that $\alpha^{(1)}$ in step 8 is not defined and the iteration cannot continue. According to [3] this is exactly a breakdown of the second kind. The case $Pu_k = Mu_k$, $P^H v_k = M^H v_k$ can be dealt with similarly. □

Note that QMR also encounters a breakdown due to $(q^{(1)})^H v^{(1)} = 0$ (cf. [4, Figure 2.8]) although this is not referred to as pivot breakdown.

Of course, since in most situations applying a preconditioner $P \neq I$ might be feasible or even necessary, and $Pu_k \neq Mu_k$, $P^H v_k \neq M^H v_k$ will most likely generically hold, this kind of breakdown appears to be no intrinsic issue. It will, however, play a role for the preconditioners proposed in the next subsection and thus we shall discuss some strategies to circumvent this issue.

Although in inner-outer eigenvalue iterations it is common to start the inner solver with zero initial vectors since they are not biased towards any eigenvector directions, one could also use nontrivial starting vectors, e.g., random ones. A more appealing idea might be to set the starting vectors equal to the approximate eigenvectors of the previous TII / TRQI iteration as $x^{(0)} = u_k$, $y^{(0)} = v_k$ since u_k, v_k may already be close to the sought solutions of the current step.

A more sophisticated approach is to using the CSBCG algorithm [3] which is intrinsically designed to deal with those pivot breakdowns. The CSBCG/LAL implementation [2] of this method appears to be the most stable choice for our purpose.

3.3 Tuned preconditioners for the simultaneous solution

Now we discuss the application of tuned preconditioners for a Petrov-Galerkin method to solve both systems in Algorithm 1 in one run. If $\mathbb{S}_k$ is a tuned preconditioner for the forward linear system, then $\mathbb{S}_k^H$ should be a tuned preconditioner for the adjoint one. Additionally, by incorporating the concept of tuning, $\mathbb{S}_k$ should satisfy

$$\mathbb{S}_k u_k = Mu_k \text{ and } \mathbb{S}_k^H v_k = M^H v_k \tag{30}$$

$$\text{or} \quad \mathbb{S}_k u_k = Au_k \text{ and } \mathbb{S}_k^H v_k = A^H v_k, \tag{31}$$

respectively. It is easy to see that this cannot be achieved with an rank-one modification of a standard preconditioner as in the one-sided case or when both system are solved one by one. Instead we propose a rank-two modification of the form

$$\mathbb{S}_k = P + a_k b_k^H + c_k d_k^H, \tag{32}$$

where $P \in \mathbb{C}^{n\times n}$ is standard preconditioner, and $a_k, b_k, c_k, d_k \in \mathbb{C}^n$ are vectors yet to be determined. For the case (30) we assume w.l.o.g. that $v_k^H M u_k = 1$. Choosing $b_k := M^H v_k$, $c_k := Mu_k$ leads to

$$Pu_k + a_k + Mu(d_k^H u_k) = Mu_k, \tag{33}$$

$$P^H v_k + (M^H v_k a_k^H) v_k + d_k = M^H v_k. \tag{34}$$

Rearranging (33) for a_k, and inserting it into (34) gives

$$\begin{aligned}(P-M)^H v_k &= M^H v_k [u_k^H (P-M)^H v_k + (u_k^H d_k)] - d_k \\ &= u_k^H [(P-M)^H v_k + d_k] M^H v_k - d_k,\end{aligned}$$

such that we may choose

$$d_k := -(P - M)^H v_k,$$

which yields

$$a_k := -(P - M)u_k + (v_k^H (P - M)u_k)u_k.$$

The complete tuned preconditioner is then the rank-two modification of the standard preconditioner P, i.e.,

$$\mathbb{S}_k = P + [Mu_k, Pu_k] \begin{bmatrix} \tau_k & -1 \\ -1 & 0 \end{bmatrix} [M^H v_k, P^H v_k]^H, \tag{35}$$

where $\tau_k := v_k^H (P + M)u_k$. It is checked easily that $\mathbb{S}_k$ and $\mathbb{S}_k^H$ satisfy (30). The applications of $\mathbb{S}_k$, $\mathbb{S}_k^H$ are, using the Sherman-Morrison-Woodbury formula [15] and $f_k := P^{-1}Mu_k$, $g_k := P^{-H}M^H v_k$, $\alpha_k := v_k^H M P^{-1} M u_k$, given by

$$\begin{aligned} \mathbb{S}_k^{-1} &= P^{-1} + [f_k, u_k] \begin{bmatrix} -\alpha_k & 0 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} g_k^H \\ v_k^H \end{bmatrix} \\ &= P^{-1} + u_k v_k^H - \frac{f_k g_k^H}{\alpha_k}, \end{aligned} \tag{36a}$$

$$\mathbb{S}_k^{-H} = P^{-H} + v_k u_k^H - \frac{g_k f_k^H}{\overline{\alpha_k}}. \tag{36b}$$

For each outer iteration two extra applications of P are required to compute the vectors f_k, g_k and apply $\mathbb{S}_k$, in particular we need one solve with P and P^H. This amounts to the same extra costs as for the tuned preconditioner for the separate solution (see Section 3.1). Note that f_k, g_k as well as the scalar α_k can be constructed and stored before the inner solver is started. For a tuned preconditioner satisfying (31) the vectors defining $\mathbb{S}_k$ become $f_k := P^{-1}Au_k$, $g_k := P^{-H}A^H v_k$, $\alpha_k = v_k^H A f_k$ and using the scaling $v_k^H A u_k = 1$. When A is singular this variant of $\mathbb{S}_k$ should only be used if a nonzero eigenvalue is sought.

4 Equivalence between two-sided Rayleigh quotient iteration and two-sided simplified Jacobi-Davidson method with preconditioned iterative solves

The two-sided Jacobi-Davidson [34, 17] (TJD) is an extension of the (one-sided) Jacobi-Davidson [30, 31] to compute eigentriples of nonnormal eigenvalue problems. A first idea of such a method was already briefly mentioned in [30]. Here we discuss the simplified TJD given in Algorithm 3 which does work with one-dimensional subspaces for the right and left eigenvector approximations.

In [17, Proposition 5.5.] it was shown (using a result for one-sided methods from [29]) that for the standard eigenvalue problem simplified TJD is equivalent to TRQI

Algorithm 3: Simplified two-sided Jacobi-Davidson method (TJD) (biorthogonal version) [34, 17]

Input : Matrix A, initial vectors $u_1,\ v_1,\ v_1^H M u_1 \neq 0$.
Output: Approximate eigentriple $(\theta_{k_{\max}},\ u_{k_{\max}},\ v_{k_{\max}})$.

for $k = 1, 2, \ldots$ **do**
- Set $\theta_k = \rho(u_k,\ v_k)$;
- Compute residuals $r_{u_k} = Au_k - \theta_k M u_k$, $r_{v_k} = A^H v_k - \overline{\theta_k} M^H v_k$;
- Test for convergence;
- Solve (approximately) $s_k \perp M^H v_k$, $t_k \perp M u_k$ from

$$\Pi_1 (A - \theta_k M) \Pi_2 s_k = -r_{u_k}, \tag{37a}$$

$$\Pi_2^H (A - \theta_k M)^H \Pi_1^H t_k = -r_{v_k}, \tag{37b}$$

  where $\Pi_1 = I - \frac{M u_k v_k^H}{v_k^H M u_k}$ and $\Pi_2 = I - \frac{u_k v_k^H M}{v_k^H M u_k}$;
- Set $u_{k+1} = (u_k + s_k)/\|u_k + s_k\|_2$, $v_{k+1} = (v_k + t_k)/\|v_k + t_k\|_2$;

when all pairs of linear systems are solved by a certain number of steps of a Petrov-Galerkin-Krylov method[1] in each outer iteration. This result however does not extend to the GEP and to preconditioned solves. In this section we show that the result can be extended for the GEP and to preconditioned TRQI and TJD if the preconditioner for TRQI is chosen in a specific way. Note that by Petrov-Galerkin method with explicitly refer to any Krylov subspace method that uses biorthonormal bases for dual Krylov subspaces and performs a Petrov-Galerkin type projection for obtaining approximate solutions for dual linear systems in the considered eigenvalue methods. Therefore, we assume that no breakdown occurs or that those are dealt with appropriately (cf. the discussion in Section 3.2.1).

In the following we neglect the index k for the outer iteration as we are only interested in the inner solves. The approximate solutions $u_{k+1},\ v_{k+1}$ of the linear systems in Algorithm 1 are denoted by $u_+,\ v_+$. Furthermore we assume w.l.o.g. that u and v satisfy the normalisation $v^H M u = 1$. Then the projections in Algorithm 3 become $\Pi_1 = I - M u v^H$ and $\Pi_2 = I - u v^H M$.

Since the operators in (37a) and (37b) are adjoint to each other, the above version of TJD allows the application of Petrov-Galerkin-Krylov methods for solving both system simultaneously. If the application of a preconditioner P is desired, it has to be projected accordingly, i.e. one has to use the projected preconditioners

$$\tilde{P} := \Pi_1 P \Pi_2, \quad \tilde{P}^H := \Pi_2^H P^H \Pi_1^H \tag{38}$$

for (37a) and (37b), respectively. Note that $\tilde{P}$, $\tilde{P}^H$ are mappings from $(M^H v)^\perp$

[1] The authors of [17] used BiCG in the theorem, which should be understood as any Petrov-Galerkin type method that produces the required bases without suffering from breakdowns (personal communication with M. E. Hochstenbach).

to $v^\perp$ and $(Mu)^\perp$ to $u^\perp$. Exploiting the inherent biorthogonality relations leads to applications of both projected preconditioners (see [24],[30])

$$\tilde{P}^\dagger = \Pi^P P^{-1}, \quad (\tilde{P}^H)^\dagger = P^{-H}(\Pi^P)^H, \tag{39}$$

where

$$\Pi^P = \left(I - \frac{P^{-1}Muv^HM}{v^HMP^{-1}Mu}\right). \tag{40}$$

Note that $\Pi_2\Pi^P = \Pi^P$. In order to ensure applicability of a Petrov-Galerkin type methods one has to use, e.g., right preconditioning for one correction equation and left preconditioning for the other one, such that the pair of preconditioned correction equations become

$$\Pi_1(A-\theta M)\Pi^P P^{-1}\tilde{s} = -r_u,\ \Pi^P P^{-1}\tilde{s} = s, \tag{41}$$

$$P^{-H}(\Pi^P)^H(A-\theta M)^H\Pi_1^H t = -P^{-H}(\Pi^P)^H r_v. \tag{42}$$

The preconditioned equations for TRQI are

$$(A-\theta M)P^{-1}\tilde{u} = Mu,\ P^{-1}\tilde{u} = u_+, \tag{43}$$

$$P^{-H}(A-\theta M)^H v_+ = P^{-H}M^H v. \tag{44}$$

Note that the other way around, i.e. using left and right preconditioning in the forward and adjoint linear system, respectively, is of course also possible as well as using both left and right preconditioning for each linear system (cf. [34] for this strategy in TJD). In the following we consider the tuned preconditioner $\mathbb{S}$ with a rank-two modification which satisfies (30) ($\mathbb{S}u = Mu$, $\mathbb{S}^H v = M^H v$) and is given by (35) with its inverse (36). In order to show the equivalence of TRQI with a tuned preconditioner and simplified TJD with a standard preconditioner we require the following Lemma. The proofs of Lemma 11 and Theorem 12 mimic the ones of [29, 35, 10, 17, 11].

Lemma 11 (Generalization of [17, Lemma 5.4], [10, Lemma 1 and Lemma 3] and [11, Lemma 5.1])**.** *Let $\Pi_1 := I - Muv^H$, $\Pi_2 := I - uv^HM$, $C := A - \theta M$, $r_u = Cu$ and $r_v = C^H v$ and $v^HMu = 1$. Let P be a standard preconditioner and let the tuned preconditioner $\mathbb{S}$ satisfy (30). Introduce the subspaces*

$$\mathcal{K}_R^{(i)} = \text{span}\{Mu, C\mathbb{S}^{-1}Mu, (C\mathbb{S}^{-1})^2Mu, \ldots, (C\mathbb{S}^{-1})^iMu\},$$

$$\mathcal{L}_R^{(i)} = \text{span}\{Mu, r_u, (\Pi_1 C\Pi^{\mathbb{S}}\mathbb{S}^{-1})r_u, \ldots, (\Pi_1 C\Pi^{\mathbb{S}}\mathbb{S}^{-1})^{i-1}r_u\},$$

$$\mathcal{M}_R^{(i)} = \text{span}\{Mu, r_u, (\Pi_1 C\Pi^{P}P^{-1})r_u, \ldots, (\Pi_1 C\Pi^{P}P^{-1})^{i-1}r_u\},$$

as well as the subspaces

$$\begin{aligned}
\mathcal{K}_L^{(i)} &= \operatorname{span}\{\mathbb{S}^{-H}M^H v, \mathbb{S}^{-H}C^H\mathbb{S}^{-H}M^H v, (\mathbb{S}^{-H}C^H)^2\mathbb{S}^{-H}M^H v, \dots \\
&\qquad \dots, (\mathbb{S}^{-H}C^H)^i\mathbb{S}^{-H}M^H v\}, \\
\mathcal{L}_L^{(i)} &= \operatorname{span}\{v, \mathbb{S}^{-H}(\Pi^{\mathbb{S}})^H r_v, (\mathbb{S}^{-H}(\Pi^{\mathbb{S}})^H C^H \Pi_1^H)\mathbb{S}^{-H}(\Pi^{\mathbb{S}})^H r_v, \dots \\
&\qquad \dots, (\mathbb{S}^{-H}(\Pi^{\mathbb{S}})^H C^H \Pi_1^H)^{i-1}\mathbb{S}^{-H}(\Pi^{\mathbb{S}})^H r_v\}, \\
\mathcal{M}_L^{(i)} &= \operatorname{span}\{v, P^{-H}(\Pi^P)^H r_v, (P^{-H}(\Pi^P)^H C^H \Pi_1^H)P^{-H}(\Pi^P)^H r_v, \dots \\
&\qquad \dots, (P^{-H}(\Pi^P)^H C^H \Pi_1^H)^{i-1}P^{-H}(\Pi^P)^H r_v\}.
\end{aligned}$$

For every $i \geq 1$ we have $\mathcal{K}_R^{(i)} = \mathcal{L}_R^{(i)} = \mathcal{M}_R^{(i)}$ and $\mathcal{K}_L^{(i)} = \mathcal{L}_L^{(i)} = \mathcal{M}_L^{(i)}$.

Proof. Using (36a) and $v^H M u = 1$ we have that $\Pi^{\mathbb{S}}\mathbb{S}^{-1} = \Pi_2\mathbb{S}^{-1} = \mathbb{S}^{-1}\Pi_1$. Then $\mathcal{K}_{\mathrm{R}}^{(i)} = \mathcal{L}_{\mathrm{R}}^{(i)}$ and $\mathcal{K}_{\mathrm{L}}^{(i)} = \mathcal{L}_{\mathrm{L}}^{(i)}$ follow directly from [17, Lemma 5.4],[11, Lemma 5.1] (applied to $C\mathbb{S}^{-1}$ and $\mathbb{S}^{-H}C^H$). Furthermore

$$\Pi^{\mathbb{S}}\mathbb{S}^{-1} = \mathbb{S}^{-1} - uv^H = P^{-1} - \frac{P^{-1}Muv^H MP^{-1}}{v^H MP^{-1}Mu} = \Pi^P P^{-1},$$

as well as $\mathbb{S}^{-H}(\Pi^{\mathbb{S}})^H = P^{-H}(\Pi^P)^H$ using (36a), (36b) and both $\mathcal{L}_{\mathrm{R}}^{(i)} = \mathcal{M}_{\mathrm{R}}^{(i)}$ and $\mathcal{L}_{\mathrm{L}}^{(i)} = \mathcal{M}_{\mathrm{L}}^{(i)}$ follow immediately. □

Theorem 12 (Generalization of [17, Proposition 5.5],[11, Theorem 5.2] and [10, Theorem 4])**.** *Let u, v be approximate right and left eigenvectors of A, M, normalized such that $v^H Mu = 1$. Moreover, $s^{(i)}$, $t^{(i)}$ denote the approximate solutions to the correction equations of simplified TJD obtained with a Petrov-Galerkin method using standard preconditioners P, P^H. Then for the approximate solutions $u^{(i+1)}$, $v^{(i+1)}$ to the TRQI equations obtained after $i+1$ steps of the same Petrov-Galerkin method applying the tuned preconditioner $\mathbb{S}$ with $\mathbb{S}u = Mu$ and $\mathbb{S}^H v = M^H v$ it holds*

$$u^{(i+1)} = \mu_1(s^{(i)} + u) \quad \text{and} \quad v^{(i+1)} = \mu_2(t^{(i)} + v),$$

for some constants μ_1, μ_2.

Proof. The spaces spanned by i steps of a preconditioned Petrov-Galerkin method applied to the JD correction equations are

$$\begin{aligned}
&\operatorname{span}\left\{r_u,\ (\Pi_1 C\Pi^P P^{-1})r_u, \dots, (\Pi_1 C\Pi^P P^{-1})^{i-1}r_u\right\}, \\
&\operatorname{span}\left\{P^{-H}(\Pi^P)^H r_v,\ (P^{-H}(\Pi^P)^H C^H \Pi_1^H)P^{-H}(\Pi^P)^H r_v, \dots\right. \\
&\qquad \left.\dots, (P^{-H}(\Pi^P)^H C^H \Pi_1^H)^{i-1}P^{-H}(\Pi^P)^H r_v\right\}
\end{aligned}$$

which, according to Lemma 11 and with $\Pi^{\mathbb{S}} = \Pi_2$, are equal to

$$\begin{aligned}
&\operatorname{span}\left\{r_u,\ (\Pi_1 C\Pi_2\mathbb{S}^{-1})r_u, \dots, (\Pi_1 C\Pi_2\mathbb{S}^{-1})^{i-1}r_u\right\}, \\
&\operatorname{span}\left\{\mathbb{S}^{-H}\Pi_2^H r_v,\ (\mathbb{S}^{-H}\Pi_2^H C^H \Pi_1^H)\mathbb{S}^{-H}\Pi_2^H r_v, \dots\right. \\
&\qquad \left.\dots, (\mathbb{S}^{-H}\Pi_2^H C^H \Pi_1^H)^{i-1}\mathbb{S}^{-H}\Pi_2^H r_v\right\},
\end{aligned}$$

respectively. Let the columns of $W^{(i)}$ and $Z^{(i)}$ span the biorthonormal bases for these two spaces generated by the Petrov-Galerkin method. It holds $Mu \perp Z^{(i)}$ and $v \perp W^{(i)}$ such that

$$u^H M^H Z^{(i)} = v^H W^{(i)} = 0, \quad \Pi_1 W^{(i)} = W^{(i)}, \ \Pi_1^H Z^{(i)} = Z^{(i)}.$$

The approximate solutions $s^{(i)}$, $t^{(i)}$ are then given by $s^{(i)} = \Pi_2 \mathbb{S}^{-1} \tilde{s}^{(i)} = \mathbb{S}^{-1} \Pi_1 \tilde{s}^{(i)}$, $\tilde{s}^{(i)} = W^{(i)} w$, $t^{(i)} = Z^{(i)} z$, where

$$\begin{aligned} w &= -(T^{(i)})^{-1} (Z^{(i)})^H r_u = -(T^{(i)})^{-1} (Z^{(i)})^H A u, \\ z &= -(T^{(i)})^{-H} (W^{(i)})^H \mathbb{S}^{-H} \Pi_2^H r_v = -(T^{(i)})^{-H} (W^{(i)})^H \mathbb{S}^{-H} A^H v \end{aligned}$$

with $T^{(i)} := (Z^{(i)})^H C \mathbb{S}^{-1} W^{(i)}$. There we have used that $\mathbb{S}^{-H} \Pi_2^H r_v = \Pi_1^H (\mathbb{S}^{-H} A^H v - \overline{\theta} \mathbb{S}^{-H} M^H v) = \Pi_1^H (\mathbb{S}^{-H} A^H v - \overline{\theta} v)$ and $\Pi_1^H v = 0$. Consequently,

$$\begin{aligned} s^{(i)} &= -\mathbb{S}^{-1} W^{(i)} (T^{(i)})^{-1} (Z^{(i)})^H A u, \\ t^{(i)} &= -Z^{(i)} (T^{(i)})^{-H} (W^{(i)})^H \mathbb{S}^{-H} A^H v. \end{aligned}$$

For the linear systems of TRQI we know by Lemma 11 that the columns of $[Mu, \ W^{(i)}]$ and $[v, \ Z^{(i)}]$ are biorthonormal bases of

$$\begin{aligned} &\operatorname{span}\left\{ Mu, C\mathbb{S}^{-1} Mu, \ldots, (C\mathbb{S}^{-1})^i Mu \right\}, \\ \operatorname{span}&\left\{ \mathbb{S}^{-H} M^H v, \mathbb{S}^{-H} C^H \mathbb{S}^{-H} M^H v, \ldots, (\mathbb{S}^{-H} C^H)^i \mathbb{S}^{-H} M^H v \right\}. \end{aligned}$$

The approximate solutions are then given by $u^{(i+1)} = \mathbb{S}^{-1} \tilde{u}^{(i+1)}$, $\tilde{u}^{(i+1)} = \mu_1 Mu + W^{(i)} p$ and $v^{(i+1)} = \mu_2 v + Z^{(i)} q$, where $\mu_1, \mu_2 \in \mathbb{C}$ and $p, q \in \mathbb{C}^i$ are determined by

$$\hat{T}^{(i)} \begin{bmatrix} \mu_1 \\ p \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad (\hat{T}^{(i)})^H \begin{bmatrix} \mu_2 \\ q \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

with

$$\hat{T}^{(i)} = \begin{bmatrix} v^H C \mathbb{S}^{-1} Mu & v^H C \mathbb{S}^{-1} W^{(i)} \\ (Z^{(i)})^H C \mathbb{S}^{-1} Mu & T^{(i)} \end{bmatrix}.$$

The $(1,1)$-entry in $\hat{T}^{(i)}$ is zero and p, q are hence obtained from

$$p = -\mu_1 (T^{(i)})^{-1} (Z^{(i)})^H A u, \quad q = -\mu_2 (T^{(i)})^{-H} (W^{(i)})^H \mathbb{S}^{-H} A^H v,$$

and hence

$$u^{(i+1)} = \mu_1 (s^{(i)} + \mathbb{S}^{-1} Mu), \quad v^{(i+1)} = \mu_2 (t^{(i)} + v),$$

from which the desired result follows. □

Corollary 13. *Lemma 11 and Theorem 12 also reveal an equivalence of TRQI and TJD in the unpreconditioned case ($P = I$) for the GEP when the tuning operator*

$$\mathbb{T} = I + [Mu, u] \begin{bmatrix} v^H u + 1 & -1 \\ -1 & 0 \end{bmatrix} [M^H v, v]^H \tag{45}$$

which satisfies $\mathbb{T}u = Mu$ and $\mathbb{T}^H v = M^H v$ is used in TRQI and TJD is necessarily preconditioned by $\Pi_1\Pi_2$ (cf. (38)). If $M = I$ these actions are not required which immediately gives [17, Proposition 5.5.].

Remark 14. *The above theorems also hold when a fixed shift is used, i.e., $\theta_k \equiv \theta \; \forall k \geq 1$, leading to an equivalence result of TII and simplified TJD.*

5 Numerical examples

We run numerical experiments regarding the discussed convergence properties of TII / TRQI in Section 2, the preconditioning and tuning strategies in Section 3, and the equivalence of TRQI and simplified TJD in Section 4. All experiments were carried out using MATLAB® 7.11.0 on a compute server using 4 Intel®Xeon®@2.67 GHz CPUs with 8 cores per CPU and 1 TB RAM.

5.1 Convergence of inexact methods

At first we verify the convergence results of Section 2 using the nuclear reactor example from [11, Example 5.1]. The dimension of this generalized eigenvalue problem is $n = 2048$ such that the occurring linear systems can be solved cheaply using the MATLAB® backslash. For the inexact solves the CSBSG/LAL [2] method was employed and no preconditioning was required. We look for the eigenvalue $\lambda = 8.0097$ and and its associated right and left eigenvectors. The shift for TII was set to $\theta = 8$. The initial vectors u_1, v_1 are the perturbed eigenvectors corresponding to λ which were generated using the `eigs(A,M,1,8)` and `eigs(A',M',1,8)` commands. The perturbation was chosen small enough such that TRQI, whose convergence strongly depends on the given initial vectors, converged to the sought eigentriple. In Figure 5.1 $\max(\|r_{u_k}\|, \|r_{v_k}\|)$ is plotted versus the outer iteration number k for both methods. As predicted by Theorem 2 inexact TII stagnates for two different fixed inner accuracies ($\xi^{\mathrm{R/L}} = 0.1,\ \ 10^{-4}$). It achieves the same convergence speed as with exact solves when decreasing inner tolerances ($\xi_k^{\mathrm{R/L}} = \min\left(0.1, 0.1\|r_{u_k/v_k}\|\right)$) are used as proposed by Theorem 3. A similar observation can be made for inexact TRQI in the right plot, although there the difference between fixed ($\xi^{\mathrm{R/L}} = 0.9$) and decreasing inner tolerances ($\xi_k^{\mathrm{R/L}} = \min\left(0.5, \|r_{u_k/v_k}\|\right)$) is only marginal due to the fast speed of convergence and the mild nature of the problem.

5.2 Preconditioned solves and tuning

For investigating the performance of the inexact solves using the proposed tuned preconditioners we use the IFISS 3.2 package [28] to discretize a convection-diffusion equa-

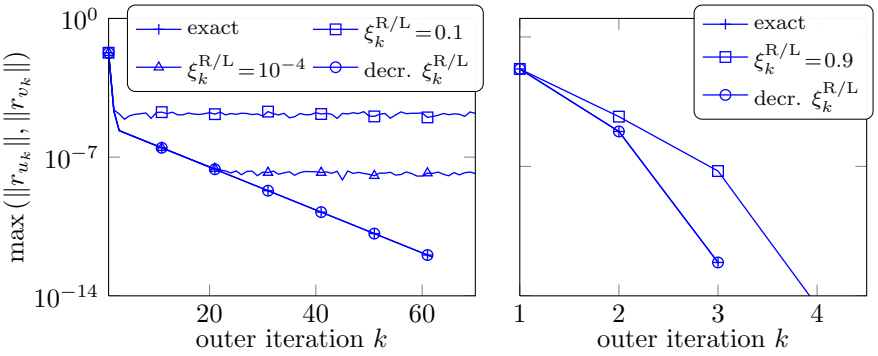

Figure 1: Convergence history of exact and inexact TII (left), TRQI (right) for the reactor example using fixed and decreasing inner tolerances. The decreasing tolerances for TII and TRQI where set via $\xi_k^{R/L} = \min\left(0.1, 0.1\|r_{u_k/v_k}\|\right)$ and $\xi_k^{R/L} = \min\left(0.5, \|r_{u_k/v_k}\|\right)$, respectively.

tion on $(-1,1)^2$ by Q1 finite elements on a uniform grid. The matrices A, M are provided by the test example T-CD2 on a 256×256 grid which gives $n = 66049$. We look for the eigentriple corresponding to the eigenvalue $\lambda = 2.4508 \cdot 10^3$. The shift for TII is $\theta = 2.5 \cdot 10^3$ and the starting vectors for TRQI were constructed as in the previous example. Both methods were terminated when $\max(\|r_{u_k}\|, \|r_{v_k}\|) < 10^{-10}$. We use decreasing inner tolerances $\xi_k^{R/L} = \min\left(\varphi, \varphi\|r_{u_k/v_k}\|\right)$ with $\varphi = 0.1$ (0.5) in TII (TRQI). The linear systems are solved separately with GMRES and simultaneously with CSBCG. Without preconditioning the inner solvers did not converge at all or within a reasonable amount of time. The standard preconditioner P is an incomplete LU decomposition of $A-\theta M$ with a drop tolerance of 0.01 and P^H was chosen for the adjoint linear system. The tuned preconditioners are $\mathbb{P}_k$, $\mathbb{Q}_k$ within GMRES, and $\mathbb{S}_k$ within CSBCG, and are given as in Section 3. We used the M-variants which satisfy

$$\mathbb{P}_k u_k = M u_k, \quad \mathbb{Q}_k v_k = M^H v_k \quad \text{and} \quad \mathbb{S}_k u_k = M u_k, \quad \mathbb{S}_k^H v_k = M^H v_k$$

(cf. (25),(28),(35)) as well as the A-variants (cf. (26),(29)) where

$$\mathbb{P}_k u_k = A u_k, \quad \mathbb{Q}_k v_k = A^H v_k \quad \text{and} \quad \mathbb{S}_k u_k = A u_k, \quad \mathbb{S}_k^H v_k = A^H v_k.$$

Table 1 gives the required outer iterations k, the total inner iterations i (i.e., matrix vector products with $A-\theta_k M$) and their average number over all outer iterations, the total number of applications with P and P^H, and the consumed CPU time for this experiment. Clearly, using tuned preconditioners leads to a decreased number of inner iterations compared to the application of the standard preconditioners in all settings except for GMRES within TRQI and the A-variant. This reduction is more significant for TII than for TRQI because of the significantly higher number of outer iterations such that savings regarding the runtime are more obvious for TII. For the majority of

Table 1: Results for inexact TII and TRQI using standard and tuned preconditioners for the IFISS example. There, k and i are the required outer and total inner iterations.

Methods	Prec.	k	i	aver.	# precs	time
TII - GMRES	standard	38	936	25	936	30.6
	tuned, M	39	325	9	401	16.4
	tuned, A	38	679	18	753	24.2
TII - CSBCG	standard	38	1436	39	1436	29.1
	tuned, M	38	856	23	930	19.9
	tuned, A	38	650	18	724	15.7
TRQI - GMRES	standard	4	174	58	174	5.9
	tuned, M	4	110	37	116	3.8
	tuned, A	4	190	63	196	5.8
TRQI - CSBCG	standard	4	258	86	258	5.2
	tuned, M	4	228	76	234	4.6
	tuned, A	4	194	65	200	4.0

cases, CSBCG required more inner iterations then GMRES, but thanks to its short recurrence formulation, does so in less time. The storing and orthogonalization of the basis vectors of GMRES is more expensive than the additional iterations, and inherent matrix vectors products, of the longer runs of CSBCG. From this one should by no means conclude that the simultaneous solution via methods such as CSBCG is in general the most efficient way. Good results for the sequential solution can also be acquired by other short recurrence methods, e.g., restarted GMRES, BiCGstab or IDR(s).

In Figure 5.2 the inner iterations are plotted against the outer iterations. The two plots on the left for inexact TII show that, as predicted by Theorem 7, the number inner iterations increases along the outer iteration when a standard preconditioners is employed. Using tuned preconditioners not only reduces the number of required inner iterations for GMRES (top left plot), but also keeps this number approximately constant after a start up time in the beginning. The effect is similar for CSBCG, although there the number of inner iterations shows a more oscillating behavior.

Although for TRQI (right plots) the reduction of the number of inner iterations is also given, this number does still increase as the outer iteration proceeds. This increase seems to be, however, smaller than for the standard preconditioner which is particularly visible in the CSBCG experiment.

To conclude, the inexact two-sided methods with preconditioned inner solves show a similar behavior as the one-sided methods as it was, e.g., investigated in [11, Theorem 3.5] using GMRES as inner solver.

Comparing the M-and A-variants of the tuned preconditioners in Table 1 and Figure 5.2, it appears that the M-variant works better within CSBCG (bottom plots) whereas GMRES (top plots) seems to benefit more from the A-variant. This was indeed also

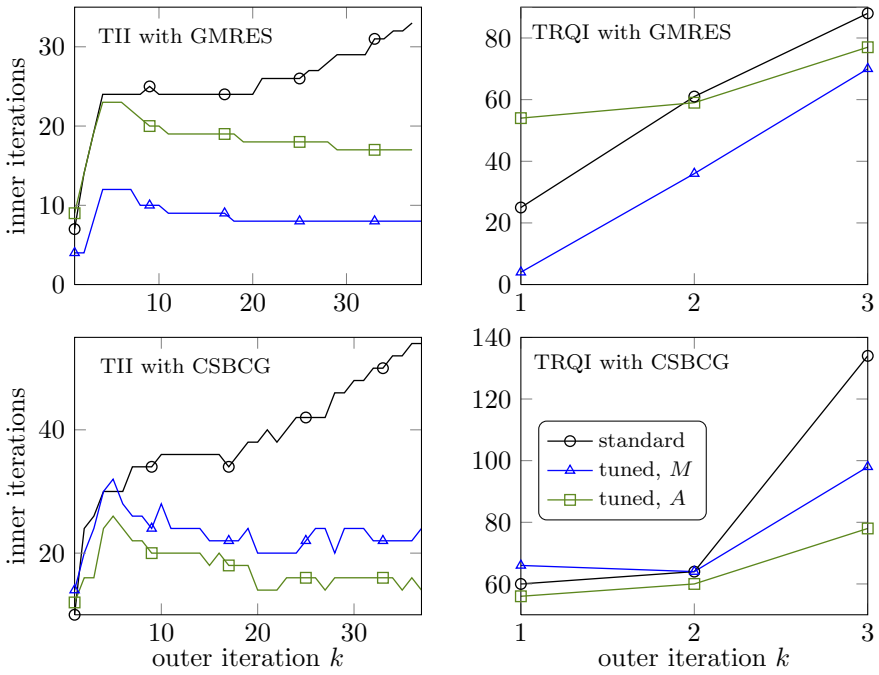

Figure 2: Progress of the required inner iterations against the outer iterations for inexact TII and TRQI using different solvers and preconditioners for the IFISS example.

observed in other experiments for TII not reported here. For TRQI slightly different choices regarding the initial vectors and the inner accuracy lead to different behavior w.r.t. the M- and A-variant of the tuned preconditioner.

5.3 Equivalence of preconditioned RQI and BiJD

We use the same example as above to investigate the equivalence of TRQI and simplified TJD which was shown in Section 4. To this end we restrict the maximum number of inner iterations to 8 (7) for TRQI (TJD) and do not stop when a certain inner accuracy is met. The short recurrence formulation of CSBCG might introduce rounding errors which could spoil the results. Hence, we employ the basic two-sided Lanczos method [26, Algorithm 7.2] with re-biorthogonalization of the generated dual Krylov bases. In line with Theorem 12 and Corollary 13 we use the standard (P, P^H) and the M-variant of the tuned ($\mathbb{S}_k$) preconditioners, as well as the tuning operator $\mathbb{T}_k$ from (45) and no preconditioner at all for the inner solves.

The convergence history for 20 outer iterations of TRQI, TJD is illustrated in Figure 5.3. In the left plot shows the results when $P = I$. As predicted by Corollary 13, TJD and TRQI are equivalent when the tuning operator $\mathbb{T}_k$ is applied to the linear

systems. Not surprisingly, this operator has no effect for TJD. Only TRQI without the application of $\mathbb{T}_k$ shows a different behavior which would be identical to the other ones if $M = I$. We also see that rounding errors induce minor differences between TRQI and TJD using $\mathbb{T}_k$ in the final outer iterations. In other similar experiments (not reported here) these differences can be larger if more inner or outer iterations are employed.

Similar observations can be made for the preconditioned case in the right plot. As proposed by Theorem 12, TJD using the standard and tuned preconditioner as well as TRQI using the tuned preconditioner give the same results. Again, only TRQI with a standard preconditioner shows a different residual history which would also be the case when $M = I$.

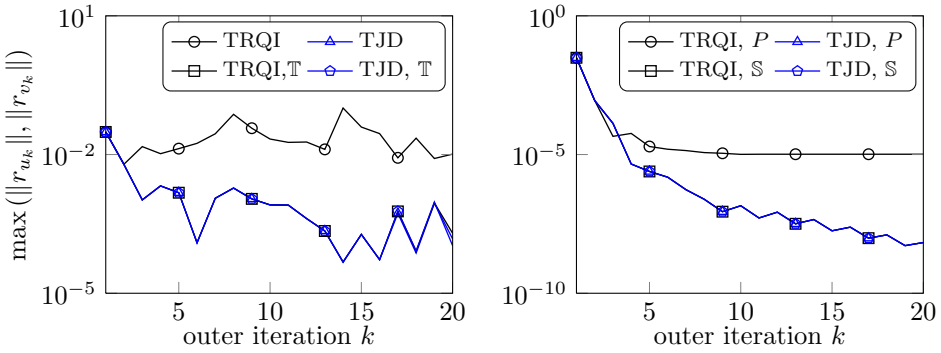

Figure 3: Convergence history of inexact TRQI and simplified TJD with 8, respectively 7, iterations of two-sided Lanczos using no preconditioner and tuning operator (left), standard and tuned preconditioner (right).

6 Conclusions

We have discussed, reviewed and extended the convergence analysis on exact and inexact two-sided inverse iteration and Rayleigh quotient iteration established in [17] to the generalized eigenvalue problem. We showed that, if inexact solves are used with a prescribed decreasing solve tolerance then the inexact two-sided methods recover the convergence rates of the exact two-sided methods, that is linear convergence for inexact two-sided inverse iteration and locally cubic convergence for inexact two-sided Rayleigh quotient iteration.

Moreover, we considered preconditioners for the inner iteration and extended the results on the tuned preconditioner for one-sided inverse iteration and Rayleigh quotient iteration [8, 7, 11] to the two-sided methods, where the forward and adjoint linear systems are solved simultaneously and therefore a rank-two modification of the standard preconditioner has to be used for the tuning strategy.

Finally we showed that the equivalence of inexact two-sided Rayleigh quotient iteration and inexact two-sided Jacobi-Davidson method (without subspace expansion) which was established in [17] for the standard eigenproblem without a preconditioner (when a certain number of steps of a Petrov-Galerkin-Krylov method is used), also holds for the generalized preconditioned eigenproblem (when a specific preconditioning strategy is applied).

Future work should validate the tuning strategies when subspace acceleration is used in TRQI and TJD as one would use in practical computations. In this case, the selection of the inner solver should be done problem-dependent.

Acknowledgements

The authors would like to thank EPSRC Network Grant EP/G01387X/1 which enabled the second author to visit the first author at the University of Bath. Moreover, we are grateful to M.E. Hochstenbach (TU Eindhoven) for various fruitful discussions.

References

[1] A. Amiraslani and P. Lancaster, *Rayleigh quotient algorithms for nonsymmetric matrix pencils*, Numerical Algorithms, 51 (2009), pp. 5–22. 10.1007/s11075-009-9286-z.

[2] R. E. Bank and T. F. Chan, *An analysis of the composite step biconjugate gradient method*, Numerische Mathematik, 66 (1993), pp. 295–319. 10.1007/BF01385699.

[3] ——, *A composite step bi-conjugate gradient algorithm for nonsymmetric linear systems*, Numerical Algorithms, 7 (1994), pp. 1–16. 10.1007/BF02141258.

[4] R. Barrett, M. Berry, T. F. Chan, J. Demmel, J. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine, and H. A. van der Vorst, *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods, 2nd Edition*, SIAM, Philadelphia, PA, 1994.

[5] J. Berns-Müller, I. G. Graham, and A. Spence, *Inexact inverse iteration for symmetric matrices*, Linear Algebra Appl., 416 (2006), pp. 389–413.

[6] R. Fletcher, *Conjugate gradient methods for indefinite systems*, in Numerical Analysis, G. Watson, ed., vol. 506 of Lecture Notes in Mathematics, Springer Berlin / Heidelberg, 1976, pp. 73–89. 10.1007/BFb0080116.

[7] M. A. Freitag and A. Spence, *A tuned preconditioner for inexact inverse iteration applied to Hermitian eigenvalue problems*, 2005. to appear in IMA J. Numer. Anal.

[8] ——, *Convergence rates for inexact inverse iteration with application to preconditioned iterative solves*, BIT, 47 (2007), pp. 27–44.

[9] ——, *Convergence theory for inexact inverse iteration applied to the generalised nonsymmetric eigenproblem*, Electron. Trans. Numer. Anal., 28 (2007), pp. 40–64.

[10] ——, *Rayleigh quotient iteration and simplified Jacobi–Davidson method with preconditioned iterative solves*, Linear Algebra and its Applications, 428 (2008), pp. 2049–2060.

[11] M. A. Freitag, A. Spence, and E. Vainikko, *Rayleigh quotient iteration and simplified Jacobi-Davidson with preconditioned iterative solves for generalised eigenvalue problems*, technical report, Dept. of Mathematical Sciences, University of Bath, 2008.

[12] R. Freund, M. Gutknecht, and N. Nachtigal, *An Implementation of the Look-Ahead Lanczos Algorithm for Non-Hermitian matrices*, SIAM Journal on Scientific Computing, 14 (1993), pp. 137–158.

[13] R. W. Freund and N. M. Nachtigal, *QMR: a quasi-minimal residual method for non-Hermitian linear systems*, Numerische Mathematik, 60 (1991), pp. 315–339. 10.1007/BF01385726.

[14] G. H. Golub, M. Stoll, and A. Wathen, *Approximation of the scattering amplitude and linear systems*, Electronic Transactions on Numerical Analysis, 31 (2008), pp. 178–203.

[15] G. H. Golub and C. F. Van Loan, *Matrix Computations (3rd ed.)*, Johns Hopkins University Press, Baltimore, MD, USA, 1996.

[16] M. R. Hestenes, *Inversion of matrices by biorthogonalization and related results*, Journal of the Society for Industrial and Applied Mathematics, 6 (1958), pp. pp. 51–90.

[17] M. E. Hochstenbach and G. L. G. Sleijpen, *Two-sided and alternating Jacobi–Davidson*, Linear Algebra and its Applications, 358(1-3) (2003), pp. 145–172.

[18] J. Lu and D. L. Darmofal, *A quasi-minimal residual method for simultaneous primal-dual solutions and superconvergent functional estimates*, SIAM J. Sci. Comput., 24 (2002), pp. 1693–1709.

[19] A. M. Ostrowski, *On the convergence of the Rayleigh quotient iteration for the computation of the characteristic roots and vectors. III (generalized Rayleigh quotient and characteristic roots with linear elementary divisors)*, Archive for Rational Mechanics and Analysis, 3 (1959), pp. 325–240.

[20] B. N. Parlett, *The Rayleigh quotient iteration and some generalizations for nonnormal matrices*, Mathematics of Computation, 28 (1974), pp. 679–693.

[21] ——, *The Symmetric Eigenvalue Problem*, vol. 20 of Classics in Applied Mathematics, SIAM, Philadelphia, PA, 1998. Corrected reprint of the 1980 original.

[22] L. Reichel and Q. Ye, *A generalized LSQR algorithm*, Numerical Linear Algebra with Applications, 15 (2008), pp. 643–660.

[23] M. Robbé, M. Sadkane, and A. Spence, *Inexact inverse subspace iteration with preconditioning applied to non-hermitian eigenvalue problems*, SIAM J. Matrix Anal. Appl., 31 (2009), pp. 92–113.

[24] J. Rommes, *Methods for eigenvalue problems with applications in model order reduction*, PhD thesis, Universiteit Utrecht, 2007.

[25] J. Rommes and G. L. G. Sleijpen, *Convergence of the dominant pole algorithm and Rayleigh quotient iteration*, SIAM Journal on Matrix Analysis and Applications, 30 (2008), pp. 346–363.

[26] Y. Saad, *Iterative Methods for Sparse Linear Systems*, SIAM, Philadelphia, PA, USA, 2003.

[27] M. A. Saunders, H. D. Simon, and E. L. Yip, *Two conjugate-gradient-type methods for unsymmetric linear equations*, SIAM Journal on Numerical Analysis, 25 (1988), pp. pp. 927–940.

[28] D. Silvester, H. Elman, and A. Ramage, *Incompressible Flow and Iterative Solver Software (IFISS) version 3.2*, May 2012. `http://www.manchester.ac.uk/ifiss/`.

[29] V. Simoncini and L. Eldén, *Inexact Rayleigh quotient-type methods for eigenvalue computations*, BIT, 42 (2002), pp. 159–182.

[30] G. L. G. Sleijpen, A. G. L. Booten, D. R. Fokkema, and H. A. van der Vorst, *Jacobi-Davidson type methods for generalized eigenproblems and polynomial eigenproblems*, BIT, 36 (1996), pp. 595–633. International Linear Algebra Year (Toulouse, 1995).

[31] G. L. G. Sleijpen and H. A. Van der Vorst, *The Jacobi–Davidson method for eigenvalue problems and its relation with accelerated inexact Newton scheme*, BIT Numerical Mathematics, 36 (1996), pp. 595–633.

[32] G. L. G. Sleijpen and H. A. van der Vorst, *A Jacobi-Davidson iteration method for linear eigenvalue problems*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 401–425.

[33] P. Smit and M. H. C. Paardekooper, *The effects of inexact solvers in algorithms for symmetric eigenvalue problems*, Lin. Alg. Appl, 287 (1999), pp. 337–357.

[34] A. Stathopoulos, *A case for a biorthogonal Jacobi–Davidson method: Restarting and correction equation*, SIAM Journal on Matrix Analysis and Applications, 24 (2002), pp. 238–259.

[35] A. Stathopoulos and Y. Saad, *Restarting techniques for the (Jacobi-)Davidson symmetric eigenvalue methods*, Electron. Trans. Numer. Anal., 7 (1998), pp. 163–181.

[36] D. Szyld and F. Xue, *Efficient Preconditioned Inner Solves For Inexact Rayleigh Quotient Iteration and their Connections to the Single-Vector Jacobi–Davidson Method*, SIAM Journal on Matrix Analysis and Applications, 32 (2011), pp. 993–1018.

[37] F. Xue and H. C. Elman, *Convergence analysis of iterative solvers in inexact Rayleigh quotient iteration*, SIAM Journal on Matrix Analysis and Applications, 31 (2009), pp. 877–899.

[38] ——, *Fast inexact subspace iteration for generalized eigenvalue problems with spectral transformation*, Linear Algebra and its Applications, 435 (2011), pp. 601–622.